



# Machine Readable Archives

CAI  
AK  
- M19

Library of Parliament



## New Acquisitions

The *Election Results, 1984* data file has been acquired from Elections Canada (formerly the Chief Electoral Office). This file provides poll by poll results from the September 1984 federal election. Data files for the federal elections in 1972, 1974, 1979 and 1980 are also available.

## EDP Records Scheduling Program

For several years the Public Archives of Canada (PAC) has depended, in large part, upon the scheduling function for the identification, appraisal, acquisition and preservation of historically valuable recorded information generated by federal government departments and agencies. Schedules are documents prepared by the records managers of federal agencies that describe the length of time groupings of related records are to be kept by an agency for operational purposes and the length of time they are to be kept by the Public Archives as dormant records in regional records centres. In order for archivists to appraise records for their archival value, agencies are required to submit these schedules to the Dominion Archivist for approval, at which time archivists are provided with an opportunity to indicate which records must be transferred to the Archives Branch of the PAC at the end of the retention periods indicated in the schedules.

The scheduling function has depended upon the establishment of a well designed records management function in each government agency. Recently, the strength of this function was reinforced by the passage of Access to Information and Privacy legislation and the issuance of a government-wide Treasury Board policy on records management. Although the scheduling function has worked reasonably well for hard-copy or textual records, it has not enhanced to any great extent information recorded in other forms or information maintained by organizational areas that were traditionally separate from the records management function. Machine readable data managed in computerized systems were rarely accounted for by the records management areas of most agencies. Although attempts to identify machine readable records held by government agencies were made in the seventies, it was not until the creation of the EDP Information Systems Section in the Machine Readable Archives (MRA) that a systematic program to identify, describe and

schedule data managed in computerized systems was developed. Recognizing the fact that individuals from several areas within an agency must be involved in the identification and scheduling of EDP records, a project team approach has been recommended. In addition to PAC representatives, a project team should be composed of the information systems representative, the records managers, the Access to Information and Privacy Coordinator, and those others who may be directly involved with the use of the data. The project team approach was tested in three pilot projects undertaken over the past year. The procedures for describing the data managed by automated systems (system overview) were tested and the results of the pilot projects were encouraging. Over the next few years it is hoped that those procedures will be used in all federal agencies.

### Schedules for EDP Records: The System Overview Approach

The development of a schedule is based on the preparation of "system overviews." An overview is designed to provide a generic picture of the information content of the data managed by the system and a large scale map of the major stages through which data flow as they are processed by the system. As a result, agencies are not required to prepare exhaustive descriptions of all of their tape and disk holdings. The intention is to identify the computerized applications supported by the agency and, based on a top-down approach, to describe the key components that can benefit from the services offered by the Public Archives.

Many systems are designed to support only one function or application (e.g., a one-time survey, issuance of labels, etc.). Others may support multiple applications, each designed to perform a unique function often based on a common pool of data (e.g., a personnel application and a financial application supported by a common database of integrated personnel and financial data). As a result, a system overview may comprise one or several application descriptions.

The elements of the system overview, other than the administrative details such as name of the institution and organization unit, are as follows:

**System Title:** A descriptive, meaningful title of the automated system.

**Application Title:** A description of the name of each application supported by the system (a system may support one or, in the case of databases, multiple applications).

**Purpose:** A description of the purpose of the application, particularly in terms of the function or related functions it is designed to perform.

**Information Content:** A summary of the information content of the data processed by the application.

**Inputs/Processes/Outputs:** An overview of the major stages through which data are processed by the application. This description is not restricted to the initial inputs and final outputs. It is based on a system or data flow chart to provide a narrative or "story" of how data are processed by the system. The frequency with which certain steps are performed is also indicated.

**Documentation/Cross References:** A description of all the key documents (and their location) that describe the specifications, procedures, codes and other supporting information related to the application.

**Software Environment:** A generic reference to the software that is used to manage the data within the application.

**Hardware Environment:** A description of the hardware used to support the application.

**System Milestones:** A summary of the major events that have affected the design of the system (production year, years in which major system conversions and changes in record formats occurred, etc.).

**User and Technical Contacts:** Names, positions and titles of the major users and individuals responsible for the technical aspects of the application.

**Subject File Cross Reference:** A reference to the appropriate title and number within the agency subject file system to which the information processed by the application is indicated. The subject file system is maintained by the agency's records manager.

**Class of Records Cross Reference:** The name and registration number of the class of records to which the information managed by the application is related. This is a requirement of the new Access to Information legislation.

**Personal Information Bank/Class Cross Reference:** The name and registration number of the Personal Information Bank (if applicable) to which the personal information managed by



## "EDP Records..."

cont'd from page 1

the application is related. This is a requirement of the new Privacy Act.

**Statement of Retention and Disposal:** The length of time that the information, regardless of the physical storage media, is required to be maintained in accordance with the legislative or policy requirements (e.g., Privacy Act, Financial Administration Act, government-wide policies and regulations, etc.) and in accordance with its operational usefulness to the agency.

**Retention and Disposal of Physical Recording Media:** Identification of the key items (e.g., process files, input forms, etc.) associated with the administration of the system; and retention periods for each of these items based on the system specifications and consultation with records managers, users and PAC staff regarding the archival value of various items, the items that should be stored in federal records centres, and the confirmation that retention periods comply with legal or policy requirements.

## Link to Archival Appraisal

The system overviews are of great benefit to the archivist in undertaking the appraisal of the records. The "purpose" statement places the system in the context of a program activity, thus

providing the archivist with the organizational information required to establish the evidential value of the data. The "information content" description provides the archivist with an opportunity to appraise the informational value of the data.

The "inputs/processes/outputs" information, by describing the steps through which the data flow as they are processed by the system application, provides the archivist with the map required to determine which form or forms of recorded information should be acquired. The "production dates" provide the archivist with an indication of the number of years of data that could be involved in a potential acquisition. The "documentation/cross references" provide the archivist with a convenient finding aid to the documentation that is required not only for a detailed understanding of the data, but also to ensure the continued accessibility of those data that are designated as having archival value. The "software and hardware environment" description provides information concerning the technical requirements of the data, and the "system milestones" will provide details on the number of years of data plus important format changes.

The "user and technical contacts" provide the archivist with the names of those individuals who have the expertise to advise on both the intellectual and technical aspects of the system and who would generally assist in the appraisal and acquisition process.

The "subject file cross reference" provides the archivist with an opportunity to perform a comprehensive appraisal of all recorded information associated with the system. The "retention and disposal" information as described in both the overview and in the related area of the subject file system provides the archivist with an indication of the length of time that various forms of recorded information associated with the system (e.g., correspondence, questionnaires, transaction files, end-of-year history files) will be retained by the agency, and an opportunity, through the appraisal process, to preserve, beyond the end of their retention periods, those data appraised as having archival value.

This approach is designed to ensure that all forms of recorded information are accounted for in one comprehensive descriptive and scheduling process. By accounting for all recorded information associated with a given system, archivists need not be concerned that they are viewing only one slice of the system. As a result, all recorded information — both hardcopy and computerized — can be viewed in their proper context within a given system.

*This article is an abridged version of a paper by John McDonald, "An Approach to the Identification and Scheduling of EDP Data," presented at the Society of American Archivists, September 2, 1984, in Washington, D.C.*

## News and Notes

### 1985 IFDO/IASSIST Conference

The 1985 IFDO/IASSIST conference was held in Amsterdam, May 20-24. Approximately one hundred forty data archivists, researchers, administrators and librarians from twenty different countries participated in the sessions. The theme of the conference, "Public Access to Public Data," provided a forum in which to discuss such problems as identifying existing data; improving access to these data; identifying barriers that prevent access; improving the dissemination of information about data files; storage media; as well as privacy issues. A list of titles of conference sessions is provided. Pre-conference abstracts of papers are available through the Steinmetz Archives in Amsterdam. Efforts are being made to publish the papers presented at the conference.

#### Session Topics:

Public Data for Policy and Research  
Overcoming Barriers to Public Access  
Archive Law: Preservation, Access,  
Responsibility  
Statistical Techniques for Preserving and  
Evaluating the Confidentiality  
of Data Releases  
Issues of Privacy and Access  
Historical Public Data  
New Media and Technology  
Comprehensive On-line Statistical  
Information Systems  
New Dimensions in Data Storage and Use  
Cooperation between Producers, Distributors  
and Users  
Varied Uses of Public Data for Policy,  
Research and Instruction  
Techniques for Secondary Analysis  
Measurement of Social Class in Public Data

Methodological Issues in Using Public Data  
The Changing Census and the Search for  
New Sources: Problems and Solutions  
The Public's Need for Data  
Longitudinal Panel Studies  
Alternative Sources of Public Data

### IASSIST '86

Next year's conference will be held in Marina Del Rey (near Los Angeles), California, May 21-25. For information about the conference contact Jackie McGee, RAND Corporation, 1700 Main St., Santa Monica, California.

The articles in this issue were prepared by Sue Gavrel, MRA.



et d'élimination comprises dans la vue d'en-  
semble du système informatique et dans la par-  
tie correspondante du système de classification

effet, la rubrique « but » situe le système dans un contexte spécifique, et l'archiviste est dès lors capable d'établir la valeur de l'information des documents. La rubrique « contenu de l'informa-  
tion » lui permet ensuite de déterminer la valeur

correspondance, d'assurer la durée de conser-  
vation des documents, d'établir les délais de transac-  
tions, recapitulations annuelles de dos-  
siers, d'assurer la durabilité de l'information, etc.

Grâce à cette méthode, toutes les catégo-  
ries d'information sont assujetties à un même

procédé de conservation qui facilite l'acquisition

de nouvelles sources : problèmes et solu-  
tions

Méthodologie propre à l'utilisation des données publiques

Le public et son besoin d'accès aux données

Enquêtes longitudinales

Autres sources de données publiques

La collection aurait lieu à

Marina Del Rey (près de Los Angeles) en

Californie, du 21 au 25 mai. Pour tous ren-

dez-vous à l'agence de l'an prochain avec

RAND Corporation, 1700 Main St., Santa

Monica, Californie.

La conférence de l'an prochain aura lieu à

## IASSIST 86

Autres sources de données publiques

Le public et son besoin d'accès aux données

de nouvelles sources : problèmes et

modifications du recensement et recherche

des données publiques

Le nouveau média et nouvelles techniques

Demandes historiques

Renewal dimensio

Nouvelles informations statistique en direct

Collaboration entre producteurs, diffuseurs

memorisation des données en matière de

Systèmes d'informations en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

La DAO.

données publiques

Techniques d'analyse secondaire dans les

Applications des données publiques et

utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

Techniques d'analyse secondaire

Techniques de politiques, recherche

et enseignement

Applications des données publiques :

et utilisateurs

et nouvelles méthodes

Information entre producteurs, diffuseurs

memorisation des données en matière de

Nouvelles dimensions en direct

**Réenvois au système de classification** : Mention du titre et du numéro du dossier dans le système de classification utilisé par l'organisme auquel sont traitées font allusion. Le système de classification des dossiers est tenu par le gestionnaire des documents de l'organisme.

*Utilisateurs et consommateurs techniques : Noms, postes et titres des principaux utilisateurs et des personnes-ressources.*

*Événements marquants : Résumé des grands systèmes qui ont influé sur la conception du système (annee de production, années où ont lieu les grandes conversions du système et les modifications majeures aux structures d'énergie, etc.).*

**Materiel :** Description du materiel utilise pour l'application.

Logiciel : Désignation générique des programmes servant à gérer les données.

Document/Renvois : Description de tous les documents-clés (et de leur emplacement) : spécifications, procédures, codes de toute autre documentation d'apprentissage.

Contenu de l'information : résumé du contenu des données traitées.

*Dur : Description du but de l'application. En parti-  
culier la fonction ou les fonctions connexes  
qui elle doit accomplir.*

**Titre de l'application :** Description de chaque application permette par le système (un système de bases de données, plusieurs applications, etc) appuyer une seule application ou, dans le cas de bases de données, plusieurs applications.

Une autre partie de l'application concerne la gestion des publications. Les publications sont regroupées en deux catégories : les publications administratives et les publications scientifiques. Les publications administratives sont destinées à la gestion interne de l'université, tandis que les publications scientifiques sont destinées à la diffusion de la recherche et de l'enseignement.

De nombreux systèmes sont construits en fonction d'une seule tâche ou application (p. ex., sondage unique, impression d'étiquette, etc.). D'autres systèmes servent parfois à des applications multiples à partir du même ensemble de fonctions dans le domaine du personnel et celui des finances (p. ex., une base de données utilisée à la fois dans les domaines de l'administration et de la finance). Par conséquent, la liste d'ensembles de systèmes peut comprendre tout ce qui précède.

L'établissement d'un calendrier de conseil-  
vation pour ce genre de documents sur  
une vue d'ensemble du système, qui sera à  
donner une idée générale du contenu de l'infor-  
matique et à fournir une représentation schéma-  
tique des grandes étapes du cheminement des  
données. Les organismes n'ont donc pas à  
électrifier leurs disques ni leurs bandes en détail.  
Ils leur suffit de faire connaitre les applications  
qui sont dans les systèmes et de décrire les éléments des  
approches pyramidale, de derrière les éléments des  
Applications qui sont utilisées pour établir profit des services des

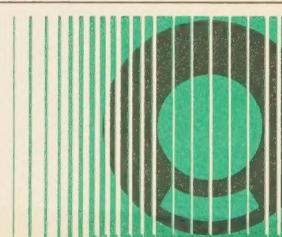
Calendriers de conservation des documents : Vue

ation à la Division des archives ordinolinguistiques (DAO) pour qu'un programme système matique le joue. Fait donne les nombreux secteurs sociaux par la question des catégories de communication, il a été décidé de privilégier le travail préparatoire. Outre les représentants des APC, on a également bon de regrouper un responsable des systèmes d'information, les gestionnaires de documents, coordinateur de l'accès à l'information, les modalités de trois projets pilotes, et enfin le cadre de trois projets pilotes. L'ensemble de ces modalités automatiques (vise des éensemblages) ont été testées, et les résultats du système ont encouragé les auteurs à ces modifications. On espère que tous les organismes fédérax se conformeront à ces modalités au cours des prochaines années.

# La gestion de la conservation des documents informatiques

#### **Nouvelles acquisitions**

# Archives ordinaires



3 1761 11550408 6